Capítulo 9

Pronóstico de Variables Climatológicas del Estado de Tlaxcala

Silvia Herrera Cortés[1]
Bulmaro Juárez Hernández[2]

Abstract: The weather forecast variables plays an important role in the activities of man by allowing anticipating adaptation measures to minimize the effects caused by climate variability. The analysis of the behavior of the climatic variables whose numerical record of their characteristics are longitudinal, can be done with time series. In this paper, the Box-Jenkins methodology is applied to analyze a database on minimum temperature, rainfall and weekly maximum five weather stations in the state of Tlaxcala.

keywords:time series, stationary processes, autocovariance function, partial autocorrelation function, prognosis.

9.1 Introduction

In September 2013, the fifth assessment report AR5 spread by the Intergovernmental Panel on Climate Change (IPCC). Among the main statements of this report, it is stated that warming of the climate system is unequivocal, and that since the 50s, the changes we are seeing are unprecedented in some time ranges ranging from decades to millennia (Figure 9.1) [IPCC2013].

[1]silvia_mat83@yahoo.com.mx. Facultad de Ciencias Físico-Matemáticas, Benemérita Universidad Autónoma de Puebla.

[2]bjuarez@fcfm.buap.mx. Facultad de Ciencias Físico-Matemáticas, Benemérita Universidad Autónoma de Puebla.

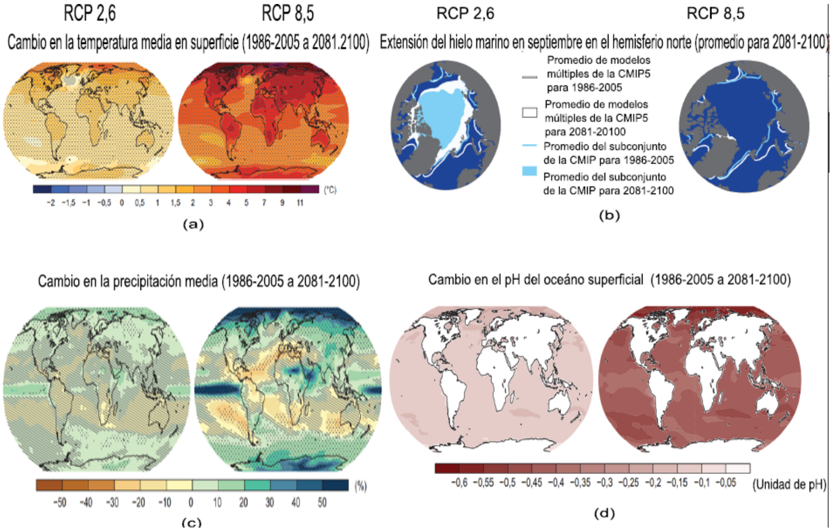

Figura 9.1: Scenarios for the change in temperature on the globe, ext sea ice, rainfall and surface ocean pH Source:[IPCC2013].

Since 2003, the research project Development developed the capabilities to stage 2 for Adaptation to Climate Change in Central America, Cuba and Mexico supported by the Programme of the United Nations Desrrollo (UNDP) and is coordinated in Mexico the National Institute of Ecology of SEMARNAT (INE-SEMARNAT). In Mexico, the previous project was conducted as a case study in the state of Tlaxcala, covering three sectors: water resources, forestry and agriculture [INE2007].

Given the above, the forecast weather variables plays an important role in the activities of man by allowing anticipating adaptation measures to minimize the effects caused by climate variability. Studying the behavior of the climatic variables whose numerical record of their characteristics are longitudinal, can be related to the study of time series data in order to forecast. The aim of this paper is to analyze time series to a database on rainfall, minimum temperature and maximum weekly temperature of five stations located in the state of Tlaxcala. For data analysis software free R is used [R]. Then basics of time series are presented.

9.2 Basics for time series

Definición 2 *If $\{X_t\}$ is a weakly stationary time series [Brokwell1987], defined*

$$\gamma_X(h) = \gamma_X(h, 0) = Cov(X_{t+h}, X_t) \quad and$$

$$\rho_X(h) = \frac{\gamma_X(h)}{\gamma_X(0)}, \;\; \forall \;\; h \in \mathbb{Z},$$

as a function autocovariance and autocorrelation function of $\{X_t\}$ respectively.

Autocovariance function of sample and sample autocorrelation is defined as

$$\widehat{\gamma}(h) = n^{-1} \sum_{t=1}^{n-|h|} \left(X_{t+|h|} - \overline{X}_n\right) \left(X_t - \overline{X}_n\right) \quad \text{and}$$

$$\widehat{\rho}(h) = \frac{\widehat{\gamma}(h)}{\widehat{\gamma}(0)}$$

respectively, with $\overline{X}_n$ the sampling process and n the sample size average.

Definición 3 *Consider the series of work, $y_b^*, y_{b+1}^*, \ldots, y_n^*$. The autocorrelation function (FAM) and the sample partial autocorrelation function (FAMP) in the phase shift k as defined*

$$r_k = \frac{\sum_{t=b}^{n-k}(y_t^* - \overline{y})(y_{t+k}^* - \overline{y^*})}{\sum_{t=b}^{n}(y_t^* - \overline{y^*})} \quad and$$

$$r_{kk} = \begin{cases} r_1, & si \quad k = 1, \\ \frac{r_k - \sum_{j=1}^{k-1} r_{k-1,j} r_{k-l}}{1 - \sum_{j=1}^{k-1} r_{k-1,j} r_j}, & si \quad k = 2, 3, \ldots, \end{cases}$$

respectively, where, $\overline{y^*} = \frac{\sum_{t=b}^{n} y_t^*}{n-b+1}$, $r_{kj} = r_{k-1,j} \; r_{k-1,k-j}$, to $j = 1, 2, \ldots, k-1$, y y_i^*, $i = b, b+1, \ldots, n$ may be the original values of the time series or transformed values [Kolcher2007].

Definición 4 *A process $ARMA(p, q)$ is a process of the form*

$$(1 - \phi_1 B - \phi_2 B^2 - \ldots - \phi_p B^p) X_t = c + (1 + \theta_1 B + \theta_2 B^2 + \ldots + \theta_q B^q) Z_t \;\; , \quad (9.2.1)$$

provided that the roots of $1 - \phi_1 z \phi_2 z^2 - \ldots - \phi_p z^p = 0$ are outside the unit circle [Guerrero1987].

The exprexión (9.2.1), when $p = 0$ a process of moving averages of order q ($MA(q)$) is obtained in the case $q = 0$, you get a autoregressive process of order p ($AR(p)$) [Guerrero1987]. To determine the model $ARMA(p, q)$ to set a database is necessary to observe the behavior of the FAM and FAMP [Kolcher2007]. The estimation of the model parameters can be performed using the least squares method or by the method of maximum likelihood cite Hamilton, in this paper, the method of maximum likelihood estimation was applied due to the properties of the estimators since the R software, used in the data analysis, this method applies.

9.3 Case Study

It has information from five weather stations in the state of Tlaxcala (Figure 9.2), for the time periods indicated in Table 9.1.

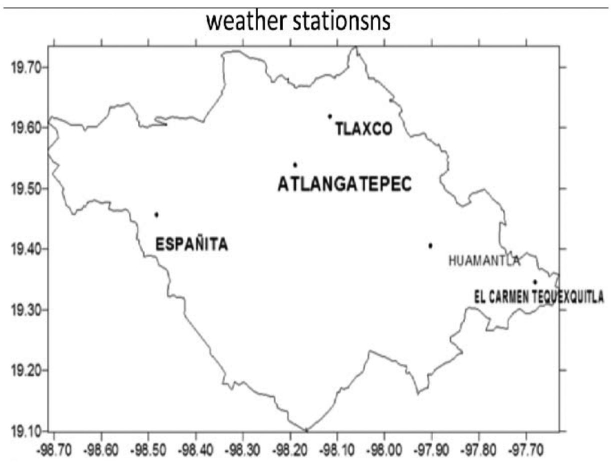

Figura 9.2: Location of the 5 weather stations studied.

The analyzed data showed quality problems related to missing data, which were completed with the daily average of 10 year record of the same month and day [Alfaro2009]. Box plots of the variables in each of the weather stations to identify seasonal cycles and information on outliers were graphed. Time series, the autocorrelation function and partial function of each of the variables studied in order to analyze the behavior of the data and identify stationarity of the data

Station	Period
Atlangatepec	1/1/1961 al 31/12/2012
Huamantla	1/1/1990 al 31/12/2012
Tlaxco	1/1/1989 al 31/12/2012
Españita	4/3/1989 al 31/12/2012
El Carmen Tequexquitla	1/1/1992 al 31/12/2012

Table 9.1: Reporting periods for 5 stations in the state of Tlaxcala.

series are plotted autocorrelation. For those data were not stationary, a transformation with first and second differences was applied and then proceeded to set the appropriate model according to the behavior of the FAM and the FAMP. To determine the best model that would fit the data, the criterion of Akaike and Ljung-Box test were followed [Kolcher2007].

In the Figure 9.3, the boxplot graphics that helped determine outliers and stationary data series for the minimum temperature, maximum and precipitation station Atlangatepec Tlaxcala are presented. In Figure 9.4, the original time series shown and the transformations applied to the data and the Ljung-Box test the temperature of Atlangatepec, Tlaxcala, Table 9.2, presents the estimation of the parameters of the best model found to fit the data observed minimum temperature Atlangatepec, Tlaxcala and Table 9.3, the observed and predicted value for the same station is presented.

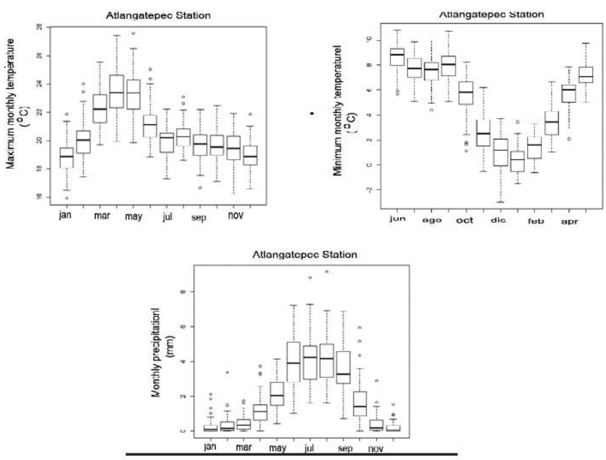

Figura 9.3: Original time series, time series and transformed Ljung-Box test for Atlangatepec station, Tlaxcala.

Following a similar procedure for the other stations, Table 9.4 shows the best

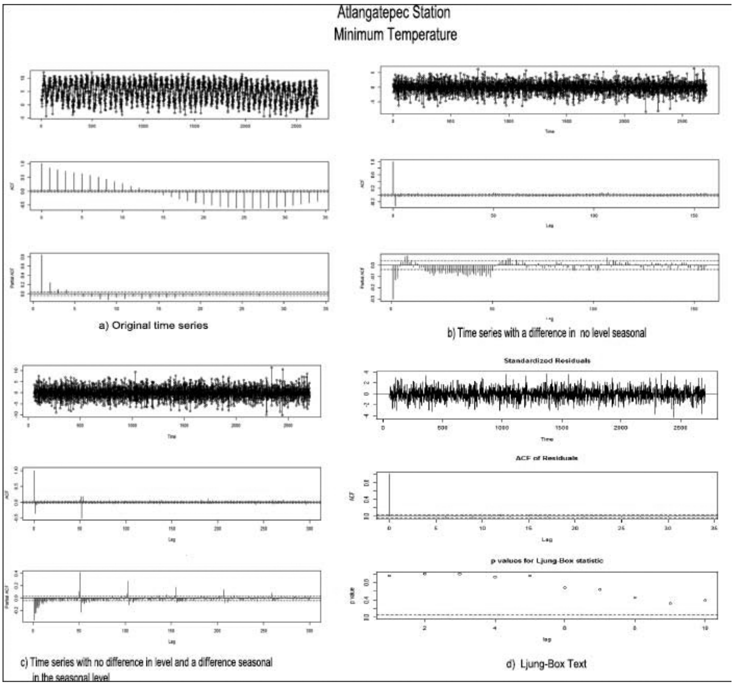

Figura 9.4: Original time series, time series and transformed Ljung-Box test for Atlangatepec station, Tlaxcala.

	ar1	ma1	ma2	sar1	sma1	sma2	sma3	intercept
	0.8001	-1.5978	0.6583	0.8001	-0.6835	-0.2049	-0.1116	-0.0001
e. s	0.2081	0.0815	0.0573	0.2082	0.0392	0.0249	0.0309	0.0001

σ^2 estimated: 4.66, log likelihood $= -5813.59$, $AIC = 11645.19$.

Table 9.2: Estimation of parameters for the model that fits the data of the minimum temperature Atlangatepec, Tlaxcala.

observed value	predicted value
0.57143	1.7215190
-0.42857	0.6916233
-0.71429	-2.9139462
-2.6429	-1.3504657
0.71429	-0.9344342
0	-0.8066220

Table 9.3: Observed and predicted value for the minimum temperature Atlangatepec, Tlaxcala.

model to fit the data found through time series and Table 9.5 it presents the observed value and the forecast time series obtained.

9.4 Conclusions

Given the importance of the variables analyzed have in agriculture state of Tlaxcala and in order to improve the climatological analysis in the study region, consider the analysis of other climatic variables such as relative humidity and solar radiation also a study using extreme theory to try to predict atypical behavior in the three climatic variables analyzed values must be performed.

In the study, boxplot graphs were useful for understanding the behavior of the database, in [ENSO] information about the years that have been classified as years Child Niña years and years Neutral to explain the presence of the observed outliers was obtained.

It is important to analyze the data with other methodologies such as neural networks or wavelet theory to compare the results.

Station	Variable	Estimating model parameters	SCE
Atlangatepec	temperature maximum	$\phi_1 = 0.9105$, $\theta_1 = -0.8679$, $\theta_2 = 0.0639$, $\theta_3 = -0.0372$ $\theta_{1,52}$=-0.7616, $\theta_{2,52} = -0.2450$, $\theta_{3,52} = 0.0066$	46.37
	precipitation	$\phi_1 = -0.4692$, $\theta_1 = -0.4343$, $\theta_{1,52} = 0.1818$, $\theta_{2,52} = -0.2706$ $\theta_{3,52} = -0.1718$	16.20
Huamatla	temperature minimum	$\theta_1 = -0.8002$, $\theta_2 = 0.3165$, $\theta_{1,52} = 0.1794$, $\theta_{2,52} = -0.4049$ $\theta_{3,52} = -0.3926$, $\theta_{4,52} = -0.1983$	4.41
	temperature maximum	$\phi_1 = 0.7900$, $\theta_1 = -0.7057$, $\theta_2 = 0.0737$, $\theta_3 = -0.3681$ $\theta_{1,52} = -0.7853$, $\theta_{2,52} = -0.2344$, $\theta_{3,52} = 0.3309$	13.48
	precipitation	$\theta_1 = -0.2745$, $\theta_2 = -0.1983$, $\theta_{1,52} = -0.7386$, $\theta_{2,52} = 0.0343$ $\theta_{3,52} = -0.2957$	22.18
Tlaxco	temperature minimum	$\theta_1 = -0.8724$, $\theta_2 = 0.3778$, $\theta_{1,52} = 0.1738$, $\theta_{2,52} = -0.4536$ $\theta_{3,52} = -0.5300$, $\theta_{4,52} = -0.1901$	35.34
	temperature maximum	$\phi_1 = -0.4894$, $\phi_2 = -0.9379$, $\theta_1 = -0.6698$, $\theta_2 = -0.1654$ $\theta_3 = -0.0097$, $\theta_{1,52} = 0.4859$, $\theta_{2,52} = 1$	19.10
	precipitation	$\theta_1 = -0.4217$, $\theta_{1,52} = -0.5616$, $\theta_{2,52} = -0.2025$, $\theta_{3,52} = -0.2359$	94.77
Españita	temperature minimum	$\phi_1 = 0.5038$, $\phi_2 = 0.1833$, $\phi_3 = 0.0661$, $\theta_1 = -1$	3.66
	temperature maximum	$\phi_1 = 0.7667$, $\phi_2 = 0.1082$, $\theta_1 = -0.8017$, $\theta_{1,52} = -0.3985$ $\theta_{2,52} = -0.3915$, $\theta_{3,52} = -0.2100$	25.43
	precipitation	$\theta_1 = 0.3307$, $\theta_{1,52} = -0.9671$, $\theta_{2,52} = 0.1553$, $\theta_{3,52} = -0.1882$	374.77
El Carmen Tequexquitla	temperature minimum	$\theta_1 = -0.5823$, $\theta_2 = 0.3064$, $\phi_{1,52} = -0.4198$, $\theta_{1,52} = 0.3348$	17.45
	temperature maximum	$\theta_1 = 0.0760$, $\theta_{1,52} = -0.761$, $\theta_{2,52} = -0.0628$	16.64
	precipitation	$\theta_1 = 0.2627$, $\theta_{1,52} = -0.6910$, $\theta_{2,52} = -0.309$	180

Table 9.4: Best model and sum of squared errors of prediction.

Station	Temperature minimum		Temperature maximum		Precipitation	
	O.V.	P.V.	O.V.	P. V.	O.V.	P.V.
Atlangatepec	0.57143	0.77192	18.143	19.776	3.1429	2.5374
	-0.42857	1.002	19.143	20.32	4.8857	2.4158
	-0.71429	1.2486	21.143	20.93	7.5714	2.2518
	-2.6429	1.5325	19	21.37	7.5714	2.0472
	0.71429	1.8677	21.571	21.739	2.5714	1.8156
	0	2.2549	21.714	22.063	4.3571	1.5992
Huamantla	4.4286	4.4282879	20.429	22.29531	6.1143	1.7113017
	1.7143	0.7434613	20.429	21.03791	0.65714	1.900622
	2.4286	2.4808470	21.286	21.40124	10.771	3.836428
	1.7143	2.5952023	21.429	18.63371	1.3143	1.479240
	4.2857	2.6581274	22.429	21.63324	1.3	10.822152
	3	2.8002909	22.286	21.20478	2.4571	0.7364535
Tlaxco	4.7857	4.9959468	20.571	23.87766	1.3429	2.6882866
	2.7143	5.3840089	20.714	22.82184	3.2	1.7364943
	2.3571	-0.6164397	22.429	22.16891	4.6857	1.7934649
	1.2143	0.9131620	21.714	23.29916	7.6714	5.4076355
	4.7857	0.4493721	23.857	23.77909	0.15714	8.8789060
	2.5	1.840886	23	24.06419	1.4714	0.3502466
Españita	7.2857	5.786217	22.714	20.95363	16.457	14.628920
	6.1429	5.764986	21	20.67117	4.8571	10.816500
	6.1429	5.736557	21.714	18.53215	9.4286	11.010197
	5.4286	5.716324	20	17.66128	8.7143	24.829147
	5.4286	5.699410	20.286	17.74430	10.286	18.830864
	4.7143	5.685203	19.714	19.30077	8.1429	7.299247
El Carmen T.	4.1429	3.3485148	19.571	21.72158	5.8571	1.2317796
	0.71429	1.1953804	20	18.65420	0	2.0623586
	-0.637	-3.5046050	20.429	22.36348	7.6429	3.1339280
	-2.199	-0.3549704	20.429	20.64627	5.4286	0.4912674
	0.857	-1.3504528	23.143	20.78707	0	2.6342668
	-0.471	-0.7751810	20.714	21.64186	0.28571	10.42004

Table 9.5: Observed and predicted value at each of the stations for the variables studied.

Bibliografía

[Brokwell1987] Brockwell J. P., Davis A. R. **Time series: Theory and Methods**.Springer-Verlag. p.p:. $8-90$. New York. 1987.

[Box-Pierce] Box-Pierce and Ljung-Box Tests. stat.ethz.ch/R-manual/R-devel/library/stats/html/box.test,html.

[Espejo2008] Espejo G. F., et al. **Cambio Climático Regional: Tendencias recientes en las series de temperatura del Pirineo central y occidental. Asociación Española de Climatología**. Serie No. 6. 2008.

[Guerrero1987] Guerrero V. (1987). **Modelos estadísticos para series de tiempo univariadas**. *V* Coloquio del departamento de matemáticas. Centro de Investigación y Estudios Avanzados del IPN.

[Hamilton1994] Douglas H. J. **Time Series Analysis**. Princeton University Press. p.p:. $1-72$ Princeton. New Jersey. 1994.

[INE2007] INE. **Proyecto: Fomento de las Capacidades para la Etapa II de Adaptación al Cambio Climático en Centroamérica, México y Cuba**. p.p:. $1-27$. 2007.

[IPCC2013] IPCC. **Climate change 2013.The physical Science Basis**. 2013.

[Alfaro2009] J. Alfaro. E., Soley J.F. (2009). **Descripción de dos métodos de rellenado de datos ausentes en series de tiempo meteorológicas. Revista de matemática: Teoría y Aplicaciones**. **16**(1). p.p: $60-75$.

[Kolcher2007] Kochler A. B. et.al. **Pronósticos, Series de tiempo y Regresión. Un enfoque aplicado.** (Traducción) Cenage Learning. Cuarta Edición. p.p.: $401-523$. 2007.

[ENSO] National Weather Service.
http://www.cpc.ncep.noaa.gov/products/precip/
CWlink/MJO/enso.shtml/history

[Peña1981] Peña S. D., Arnáiz T. G. **Criterios de selección de modelos ARIMA**. Trabajos de Estadística y de Investigación Operativa. **32**(1).p.p.:$70-93$. 1981.

[R] **R project for statistical computing**. www.r-project.org/web/packages/tseries/index.html.